2. Theorem 1.7.15 is just a restatement in terms of linear morphisms of a fact that you might have come across before: every $m \times n$ matrix can be row-reduced to reduced echelon form using row operations. Moreover, if we allow 'column operations', then any $m \times n$ matrix can be row/column-reduced to a matrix of the form appearing in Theorem 1.7.15.

   This requires the use of *elementary (row-operation) matrices* and we will investigate this result during discussion.

3. Corollary 1.7.17 allows us to provide a classification of $m \times n$ matrices based on their rank: namely, we can say that A and B are *equivalent* if there exists $Q \in \mathrm{GL}_m(\mathbb{K})$, $P \in \mathrm{GL}_n(\mathbb{K})$ such that

$$B = Q^{-1}AP.$$

   Then, this notion of equivalence defines an *equivalence relation* on $Mat_{m,n}(\mathbb{K})$. Hence, we can partition $Mat_{m,n}(\mathbb{K})$ into dictinct equivalence classes. Corollary 1.7.17 says that the equivalence classes can be labelled by the rank of the matrices that each class contains.

1.8 Dual Spaces (non-examinable)

In this section we are going to try and understand a *'coordinate-free'* approach to solving systems of linear equations and to prove some basic results on row-reduction; in particularm we will prove that 'row-reduction works'. This uses the notion of the *dual space* of a $\mathbb{K}$-vector space V. We will also see the dual space appear when we are discussing bilinear forms and the *adjoint* of a linear morphism (Chapter 3).

Definition 1.8.1. Let V be a $\mathbb{K}$-vector space. We define the *dual space of* V, denoted V^*, to be the $\mathbb{K}$-vector space

$$V^* \stackrel{def}{=} \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K}).$$

Therefore, vectors in V^* are $\mathbb{K}$-linear morphisms from V to $\mathbb{K}$; we will call such a linear morphism a **linear form on** V.

Notice that $\dim_{\mathbb{K}} V^* = \dim_{\mathbb{K}} V$.

Example 1.8.2. Let V be a $\mathbb{K}$-vector space, $\mathcal{B} = (b_1, \dots, b_n)$ an ordered basis of V. Then, for each $i = 1, \dots, n$, we define $b_i^* \in V^*$ to be the linear morphism defined as follows: since $\mathcal{B}$ is a basis we know that for every $v \in V$ we can write a unique expression

$$v = c_1 b_1 + \dots + c_n b_n.$$

Then, define

$$b_i^*(v) = c_i,$$

so that b_i^* is the function that 'picks out' the i^{th} entry in the $\mathcal{B}$-coordinate vector $[v]_{\mathcal{B}}$ of v.

Proposition 1.8.3. *Let V be a $\mathbb{K}$-vector space, $\mathcal{B} = (b_1, \dots, b_n)$ an ordered basis of V. Then, the function*

$$\theta_{\mathcal{B}} : V \to V^* \ ; \ v = c_1 b_1 + \dots + c_n b_n \mapsto c_1 b_1^* + \dots + c_n b_n^*,$$

is a bijective $\mathbb{K}$-linear morphism. Moreover, $\mathcal{B}^* = (b_1^*, \dots, b_n^*)$ is a basis of V^* called the dual basis of *$\mathcal{B}$.*

Proof: Linearity is left as an exercise to the reader.

To show that $\theta_{\mathcal{B}}$ is bijective it suffices to show that $\theta_{\mathcal{B}}$ is injective, by Theorem 1.7.4. Hence, we will show that $\ker \theta_{\mathcal{B}} = \{0_V\}$: let $v \in \ker \theta_{\mathcal{B}}$ and suppose that

$$v = c_1 b_1 + \dots + c_n b_n.$$

Then,

$$0_{V^*} = \theta_{\mathcal{B}}(v) = c_1 b_1^* + \dots + c_n b_n^* \in V^*.$$

Hence, since this is an equality of morphisms, we see that evaluating both sides of this equality at b_i, and using the defintion of b_k^*, we have

$$0 = 0_{V^*}(b_i) = (c_1 b_1^* + ... + c_n b_n^*)(b_i) = c_1 b_1^*(b_i) + ... + c_n b_n^*(b_i) = c_i, \quad \text{for every } i,$$

so that $c_1 = ... = c_n = 0 \in \mathbb{K}$. Hence, $v = 0$ and the result follows. □

Definition 1.8.4. Let $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ be a linear morphism between $\mathbb{K}$-vector spaces V, W. Then, we define the *dual of* f, denoted f^*, to be the function

$$f^* : W^* \to V^* \; ; \; \alpha \mapsto f^*(\alpha) = \alpha \circ f.$$

Remark 1.8.5. 1. Let's clarify just exactly what f^* is, for a given $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$: we have defined f^* as a function whose inputs are linear morphisms $\alpha : W \to \mathbb{K}$ and whose output is the linear morphism

$$f^*(\alpha) = \alpha \circ f : V \to W \to \mathbb{K} \; ; \; v \mapsto \alpha(f(v)).$$

Since the composition of linear morphisms is again a linear morphism we see that f^* is a well-defined function

$$f^* : W^* \to V^*.$$

We say that f^* **pulls back forms on** W **to forms on** V**.** Moreover, the function f^* is actually a linear morphism, so that $f^* \in \mathrm{Hom}_{\mathbb{K}}(W^*, V^*)$.[33]

2. Dual spaces/morphisms can be very confusing at first. It might help you to remember the following diagram

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, dashed, "f^*(\alpha)"'] & W \arrow[d, "\alpha"] \\ & \mathbb{K} \end{tikzcd}$$

It now becomes clear why we say that f^* pulls back forms on W to forms on V.

3. The $(-)^*$ operation satisfies the following properties, which can be easily checked:

   - for $f, g \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ we have $(f + g)^* = f^* + g^* \in \mathrm{Hom}_{\mathbb{K}}(W^*, V^*)$; for $\lambda \in \mathbb{K}$ we have $(\lambda f)^* = \lambda f^* \in \mathrm{Hom}_{\mathbb{K}}(W^*, V^*)$,
   - if $f \in \mathrm{Hom}_{\mathbb{K}}(V, W), g \in \mathrm{Hom}_{\mathbb{K}}(W, X)$, then $(g \circ f)^* = f^* \circ g^* \in \mathrm{Hom}_{\mathbb{K}}(X^*, V^*)$; $\mathrm{id}_V^* = \mathrm{id}_{V^*} \in \mathrm{End}_{\mathbb{K}}(V^*)$.

4. Let $\mathcal{B} = (b_1, \dots, b_n) \subset V, \mathcal{C} = (c_1, \dots, c_m) \subset W$ be ordered bases and $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Let $\mathcal{B}^*, \mathcal{C}^*$ be the dual bases of $\mathcal{B}$ and $\mathcal{C}$. Then, we have that the matrix of f^* with respect to $\mathcal{C}^*, \mathcal{B}^*$ is

$$[f^*]_{\mathcal{C}^*}^{\mathcal{B}^*} = [[f^*(c_1^*)]_{\mathcal{B}^*} \; \cdots \; [f^*(c_m^*)]_{\mathcal{B}^*}],$$

an $n \times m$ matrix.

Now, for each i, we have

$$f^*(c_i^*) = \lambda_{1i} b_1^* + ... + \lambda_{ni} b_n^*,$$

so that

$$[f^*(c_i^*)]_{\mathcal{B}^*} = \begin{bmatrix} \lambda_{1i} \\ \vdots \\ \lambda_{ni} \end{bmatrix}, \quad \text{and} \quad \lambda_{ki} = f^*(c_i^*)(b_k) = c_i^*(f(b_k)).$$

As $c_i^*(f(b_k))$ is the i^{th} entry in the $\mathcal{C}$-coordinate vector of $f(b_k)$, we see that the ik^{th} entry of $[f]_{\mathcal{B}}^{\mathcal{C}}$ is λ_{ki}, which is the ki^{th} entry of $[f^*]_{\mathcal{C}^*}^{\mathcal{B}^*}$. Hence, we have, if $A = [f]_{\mathcal{B}}^{\mathcal{C}}$, then

$$[f^*]_{\mathcal{C}^*}^{\mathcal{B}^*} = A^t.$$

[33] Check this.

Lemma 1.8.6. *Let V, W be finite dimensional $\mathbb{K}$-vector spaces, $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Then,*

- *f is injective if and only if f^* is surjective.*

- *f is surjective if and only if f^* is injective.*

- *f is bijective if and only if f^* is bijective.*

Proof: The last statement is a consequence of the first two.

($\Rightarrow$) Suppose that f is injective, so that $\ker f = \{0_V\}$. Then, let $\beta \in V^*$ be a linear form on V, we want to find a linear form α on W such that $f^*(\beta) = \alpha$. Let $\mathcal{B} = (b_1, \dots, b_n)$ be an ordered basis of V, $\mathcal{B}^*$ the dual basis of V^*. Then, since f is injective, we must have that $f(\mathcal{B}) = (f(b_1), \dots, f(b_n))$ is a linearly independent subset of W[34]. Extend this to a basis $\mathcal{C} = (f(b_1, \dots, f(b_n), c_{n+1}, \dots, c_m)$ of W and consider the dual basis $\mathcal{C}^*$ of W^*.

In terms of the dual basis $\mathcal{B}^*$ we have

$$\beta = \lambda_1 b_1^* + \dots + \lambda_n b_n^* \in V^*.$$

Consider

$$\alpha = \lambda_1 f(b_1)^* + \dots \lambda_n f(b_n)^* + 0c_{n+1}^* + \dots + 0c_m^* \in W^*.$$

Then, we claim that $f^*(\alpha) = \beta$. To show this we must show that $f^*(\alpha)(v) = \beta(v)$, for every $v \in V$ (since $f^*(\beta), \alpha$ are both functions with domain V). We use the result (proved in Short Homework 4): if $f(b_i) = g(b_i)$, for each i, with f, g linear morphisms with domain V, then $f = g$. So, we see that

$$\begin{aligned} f^*(\alpha)(b_i) &= \lambda_1 f^*(f(b_1)^*)(b_i) + \dots + \lambda_n f^*(f(b_n)^*)(b_i) + 0_V, \quad \text{using linearity of } f^*, \\ &= \lambda_1 f(b_1)^*(f(b_i)) + \dots + \lambda_n f(b_n)^*(f(b_i)) = \lambda_i, \quad \text{since } f^* \text{ pulls back forms.} \end{aligned}$$

Then, it is easy to see that $\beta(b_i) = \lambda_i$, for each i. Hence, we must have $f^*(\alpha) = \beta$.

The remaining properties are left to the reader. In each case you will necessarily have to use some bases of V and W and their dual bases. □

Remark 1.8.7.

1. Lemma 1.8.6 allows us to try and show properties of a morphism by showing the 'dual' property of its dual morphism. You will notice in the proof that we had to make a choice of a basis of V and W and that this choice was arbitrary: for a general $\mathbb{K}$-vector space there is no 'canonical' choice of a basis. In fact, every proof of Lemma 1.8.6 must make use of a basis - there is no way that we can obtain these results without choosing a basis at some point. This is slightly annoying as this means there is no 'God-given' way to prove these statements, all such attempts must use some arbitrary choice of a basis.

2. Lemma 1.8.6 does **NOT** hold for infinite dimensional vector spaces. In fact, in the infinite dimensional case it is not true that V is isomorphic to V^*: the best we can do is show that there is an injective morphism $V \to V^*$. This a subtle and often forgotten fact.

In light of these remarks you should start to think that the passage from a vector space to its dual can cause problems because there is no 'God-given' way to choose a basis of V. However, these problems disappear if we *dualise twice*.

Theorem 1.8.8. *Let V be a finite dimensional $\mathbb{K}$-vector space. Then, there is a 'canonical isomorphism'*

$$\mathrm{ev} : V \to (V^*)^* \ ; \ v \mapsto (\mathrm{ev}_v : \alpha \mapsto \alpha(v))$$

Before we give the proof of this fact we will make clear what the function ev does: since V^* is a $\mathbb{K}$-vector space we can take its dual, to obtain $(V^*)^* \stackrel{def}{=} V^{**}$. Therefore, ev_v must be a linear form on V^*, so must take 'linear forms on V' as inputs, and give an output which is some value in $\mathbb{K}$. Given the linear form $\alpha \in V^*$, the output of $\mathrm{ev}_v(\alpha)$ is $\alpha(v)$, so we are 'evaluating α at v'.

[34] You have already showed this in Short Homework 3.

The reason we say that this isomorphism is 'canonical' is due to the fact that we did not need to use a basis to define ev - the same function ev works for *any* vector space V, so can be thought of as being 'God-given' or 'canonical' (there is no arbitrariness creeping in here).

Proof: ev is injective: suppose that $\mathrm{ev}_v = 0_{V^{**}}$, so that ev_v is the zero linear form on V^*. If $v \neq 0_V$ then we can extend the (linearly independent) set $\{v\}$ to a basis of V (simply take a maximal linearly independent subset of V that contains v). Then, $v^* \in V^*$ is the linear form that 'picks out the v-coefficient' of an arbitrary vector $u \in V$ when written as a linear combination using the basis containing v. Then, we must have

$$0 = ev_v(v^*) = v^*(v) = 1,$$

which is absurd. Hence, we can't have that $v \neq 0_V$ so that ev is injective.

Hence, since $\dim V = \dim V^* = \dim V^{**}$ we see that ev is an isomorphism. □

1.8.1 Coordinate-free systems of equations or Why row-reduction works

We know that a system of m linear equations in n variables is the same thing as a matrix equation

$$A\underline{x} = \underline{b},$$

where A is the coefficent matrix of the system and $\underline{x}$ is the vector of variables. We are going to try and consider systems of linear equations using linear forms.

Let $\mathcal{S}^{(n)} = (e_1 \dots , e_n)$ be the standard basis of $\mathbb{K}^n$, $\mathcal{S}^{(n),*}$ the dual basis. Then, if α is a linear form on $\mathbb{K}^n$ we see that

$$\alpha = \lambda_1 e_1^* + \dots + \lambda_n e_n^*.$$

Hence, if $\underline{x} = x_1 e_1 + \dots + x_n e_n \in \mathbb{K}^n$ then

$$\alpha(\underline{x}) = 0 \iff \lambda_1 x_1 + \dots + \lambda_n x_n = 0.$$

Hence, $\ker \alpha = \{\underline{x} \in \mathbb{K}^n \mid \lambda_1 x_1 + \dots + \lambda_n x_n = 0\}$.

Now, suppose that $\alpha_1, \alpha_2, \dots , \alpha_m \in \mathbb{K}^{n\,*}$ are linear forms. Then,

$$\bigcap_{i=1}^{m} \ker \alpha_i = \{\underline{x} \in \mathbb{K}^n \mid \alpha_i(\underline{x}) = 0, \text{ for every } i\} .$$

So, if $\alpha_i = \lambda_{i1} e_1^* + \dots + \lambda_{in} e_n^*$, then

$$\bigcap_{i=1}^{m} \ker \alpha_i = \left\{ \underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n \mid \begin{array}{cc} \lambda_{11} x_1 + \dots + \lambda_{1n} x_n & = 0 \\ \vdots & \vdots \\ \lambda_{m1} x_1 + \dots + \lambda_{mn} x_n & = 0 \end{array} \right\}.$$

This is precisely the solution set of the matrix equation

$$A\underline{x} = \underline{0},$$

where $A = [\lambda_{ij}]$. Hence, we have translated our 'systems of linear equations' problem into one involving linear forms: namely, we want to try and understand $\bigcap_i \ker \alpha_i$, for some linear forms $\alpha_i \in \mathbb{K}^{n\,*}$.

Now, how can we interpret elementary row operations in this new framework? Of course, swapping rows is the same as just reordering the forms α_i. What happens if we scale a row by $\lambda \in \mathbb{K}$? This is the same as considering the linear form $\lambda\alpha \in \mathbb{K}^{n\,*}$. Similarly, adding row i to row j is the same as adding α_i to α_j to obtain the linear form $\alpha_i + \alpha_j$. In summary, performing elementary row operations is the same as forming linear combinations of the linear forms α_i.

The whole reason we row-reduce a matrix A to a reduced echelon form U is because the solution sets of $A\underline{x} = \underline{0}$ and $U\underline{x} = \underline{0}$ are the same (a fact we will prove shortly), and it is easier to determine solutions for the matrix equation defined by U. Since we obtain U by applying elementary row operations to A, this is the same as doing calculations in $\mathrm{span}_{\mathbb{K}}\{\alpha_i\} \subset \mathbb{K}^{n\,*}$, by what we have discussed above. Moreover, since U is in reduced echelon form this means that the rows of U are linearly independent (this is easy to see, by the definition of reduced echelon form) and because elementary row operations correspond to forming linear combinations of linear forms, we have that the linear forms that correspond to the rows of U must form a basis of the subspace $\mathrm{span}_{\mathbb{K}}\{\alpha_i\} \subset \mathbb{K}^{n\,*}$.

Definition 1.8.9. Let V be a finite dimensional $\mathbb{K}$-vector space, V^* its dual space. Let $U \subset V$ be a subspace of V and $X \subset V^*$ a subspace of V^*. We define

$$\mathrm{ann}_{V^*} U = \{\alpha \in V^* \mid \alpha(u) = 0, \text{ for every } u \in U\}, \text{ and}$$

$$\mathrm{ann}_V X = \{v \in V \mid \mathrm{ev}_v(\alpha) = 0, \text{ for every } \alpha \in X\},$$

the *annihilators of U (resp. X) in V^* (resp. V).* They are subspaces of V^* (resp. V), for any U (resp. X).

Proposition 1.8.10. *Let V be a $\mathbb{K}$-vector space, $U \subset V$ a subspace. Then,*

$$\dim V = \dim U + \dim \mathrm{ann}_{V^*} U.$$

Proof: Let $\mathcal{A} = (a_1, \dots, a_k)$ be a basis of U and extend to a basis $\mathcal{B} = (a_1, \dots, a_k, a_{k+1}, \dots, a_n)$ of V. Then, it is easy to see that $a_{k+1}^*, \dots, a_n^* \in \mathrm{ann}_{V^*} U$. Moreover, if $\alpha \in \mathrm{ann}_{V^*} U$ then

$$\alpha = \lambda_1 a_1^* + \dots + \lambda_n a_n^*,$$

and we must have, for every $i = 1, \dots, k$, that

$$0 = \alpha(a_i) = \lambda_i.$$

Hence, $\alpha \in \mathrm{span}_{\mathbb{K}}\{a_{k+1}^*, \dots, a_n^*\}$ implying that $(a_{k+1}^*, \dots, a_n^*)$ is a basis of $\mathrm{ann}_{V^*} U$. The result now follows. □

Corollary 1.8.11. *Let $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ be a linear morphism between finite dimensional vector spaces. Suppose that $A = [f]_{\mathcal{B}}^{\mathcal{C}} = [a_{ij}]$ is the matrix of f with respect to the bases $\mathcal{B} = (b_i) \subset V, \mathcal{C} = (c_j) \subset W$. Then,*

$$\mathrm{rank}\, f = \textit{ max. no. of linearly ind't columns of } A \;=\; \textit{max. no. of linearly ind't rows of } A.$$

Proof: The first equality was obtained in Lemma 1.7.14. The maximal number of linearly independent rows is equal to $\dim \mathrm{span}_{\mathbb{K}}\{\alpha_i\}$, where

$$\alpha_i = a_{i1} b_1^* + \dots + a_{in} b_n^* \in V^*.$$

Now, we have that

$$\mathrm{ann}_V \mathrm{span}_{\mathbb{K}}\{\alpha_i\} = \{v \in V \mid \mathrm{ev}_v(\alpha_i) = 0, \text{ for every } i\} = \{v \in V \mid \alpha_i(v) = 0, \text{ for every } i\},$$

and this last set is nothing other than $\ker f$.[35] Thus, by the Rank Theorem (Theorem 1.7.13) we have

$$\dim V = \dim \mathrm{im} f + \dim \ker f = \mathrm{rank}\, f + \dim \mathrm{ann}_V \mathrm{span}_{\mathbb{K}}\{\alpha_i\} = \mathrm{rank}\, f + (\dim V - \dim \mathrm{span}_{\mathbb{K}}\{\alpha_i\}),$$

using Proposition 1.8.10 in the last equality. Hence, we find

$$\mathrm{rank}\, f = \dim \mathrm{span}_{\mathbb{K}}\{\alpha_i\},$$

which is what we wanted to show. □

Proposition 1.8.12 (Row-reduction works)**.** *Let $A \in Mat_{m,n}(\mathbb{K})$, U be its reduced echelon form. Then, $\underline{x}$ satisfies $U\underline{x} = \underline{0}$ if and only if $A\underline{x} = \underline{0}$.*

Proof: Let $\alpha_1, \dots, \alpha_m \in \mathbb{K}^{n\,*}$ be the linear forms corresponding to the rows of A, $\beta_1, \dots, \beta_r$ be the linear forms corresponding to the (nonzero) rows of U (we have just seen that $r = \mathrm{rank}\, A$). Then, by our discussions above, we know (β_j) is a basis of $W = \mathrm{span}_{\mathbb{K}}\{\alpha_i\}_{i=1}^m \subset \mathbb{K}^{n\,*}$. In particular, $\mathrm{span}_{\mathbb{K}}\{\beta_j\} = W$. Now, we have

$$\mathrm{ann}_{\mathbb{K}^n} W = \{\underline{x} \in \mathbb{K}^n \mid \alpha_i(\underline{x}) = 0, \text{ for every } i\} = \{\underline{x} \in \mathbb{K}^n \mid \beta_j(\underline{x}) = 0, \text{ for every } j\}.$$

The result follows from this equality of sets since this common set is the solution set of $A\underline{x} = \underline{0}$ and $U\underline{x} = \underline{0}$. □

[35] Think about this!